Parameter spaces for stationary DGPs in spatial econometric modelling

Matthias Koch
Institute for Economic Geography and GIScience, Vienna University of Economics and Business, Nordbergstrasse 15/4/A, 1090 Vienna, Austria

Abstract

Unlike the time series literature the spatial econometric literature has not really dealt with the issue of the parameter space. This paper shows that current parameter space concepts for spatial econometric DGPs are inadequate. It proves that the parameter space proposed by Kelejian and Prucha 2008 can result in nonstationary DGPs, while the parameter space proposed by Lee and Liu 2010 can be too restrictive in applied cases. Furthermore it is discussed that the practice of row standardizing lacks a mathematical foundation.

Due to these problems concerning the current parameter space consepts, this paper provides a new definition for the spatial econometric parameter space. It is able to show which assumptions are necessary to give row standardizing the needed mathematical foundation. Finally two additional applications for the new parameter space definition concerning models with group interaction and panels with fixed cross section sample size are provided. Both applications result in parameter spaces that are substantially larger than the ones the literature would so far considered to be stationary.

1 Introduction

Spatial econometric models are widely used for empirical problems containing spatial autocorrelation. The key element to deal with spatial autocorrelation is to incorporate spatial lags into the regression model. This is similar to the case of time series where time lags are added. For the instance of linear problems in time series it is well known which parameter space configurations yield stationary data generating processes (DGPs) [see for example Hamilton 1996, chapter 1]. There is no doubt that the (admissible) parameter space for the time lag given a stationary first order autoregressive process is $(-1, 1)$. The spatial econometric literature, so far, has not dealt with the issue of the spatial parameter space. That is even reflected by the nomenclature like, in the time series literature it is common to distinguish between admissible, stationary... parameter spaces, while in spatial econometrics virtually no distinction takes place. Due to this disregard of the spatial parameter space by the literature one might not be surprised that this paper can address the following three central issues: *First* it shows that common parameter space concepts are inadequate. *Second* it proposes a new mathematical parameter space definition and *third* shows some applications for this new parameter space definition. In order to keep the nomenclature simple this paper will refer to sets containing the spatial autocorrelation parameters which result into stationary DGPs simply as spatial parameter spaces, or if it is clear form the context parameter space.

In applied spatial econometrics the following spatial parameter space concept is quiet common and this paper will refer to it as the "practitioners" approach: The approach is characterized by normalizing the spatial structure represented by the spatial weight matrix $\mathbf{W}_n$ either by dividing $\mathbf{W}_n$ with it's maximum absolute row sum (maximum row standardizing) or dividing each row of $\mathbf{W}_n$ by its absolute row sum (row standardizing) and then, assumes that the process will be stationary for spatial lag parameters in the set $(-1, 1)$. This procedure seems plausible due to the similarity to the time series approach. The "practitioners" approach of row standardizing can be seen as an approximation of the Kelejian and Prucha 2008 parameter space. Kelejian and Prucha argue that a parameter space for stationary DGPs must be in a subset of $(-1/|\tau_{\min}|, 1/|\tau_{\max}|)$, where $\tau_{\max}$ and $\tau_{\min}$ represent the maximum and minimum eigenvalue[1] of the spatial weight matrix $\mathbf{W}_n$. On the other hand Lee and Liu 2010 motivate their spatial parameter space concept an other way. Theirs is based on the Neumannseries, which states that if a matrixnorm of $\rho\mathbf{W}_n$ is smaller one that the inverse of $(\mathbf{I}_n - \rho\mathbf{W}_n)$ exists, where $\mathbf{I}_n$ is a identity matrix of size n. Additionally if the Manhattan and Infinity norms of $\rho\mathbf{W}_n$ are assumed to be smaller one, due to the Neumannseries and the subadditivity of matrix norms, the inverse of $(\mathbf{I}_n - \rho\mathbf{W}_n)$ will be bounded in row and column sums in absolute value[2]. These concepts and their background will be introduced in the

[1] Since the simple normalzing procedures like maximum row standardizing can be seen as approximation for the biggest absolute eigenvalue, some authors [ref,ref] correctly pointed out that this may result in a too restrictive parameterspace, since $(-1, 1)$ may only a subset of $1/\max_i \left\{\sum_{j=1}^{n} |w_{i,j}|\right\} (-1/|\tau_{\min}|, 1/|\tau_{\max}|)$

[2] For a mathmatical proof, see Lemma YY in the appendix

next sections in more detail.

This paper shows basically three problems are existing concerning the previous mentioned parameter spaces. *First* the parameter space concept of Kelejian and Prucha 2008, which is only considering the eigenvalues can result in nonstationary DGPs. *Second* applying the Lee and Liu 2010 parameter space and the "practitioners" approach of row standardizing on the same $\mathbf{W}_n$ can result in different parameter spaces[3]. While Lee and Liu 2010, given a spatial one forward on behind structure of $\mathbf{W}_n$ would only consider the spatial parameters to be in $(-2/3, 2/3)$, the "practitioners" approach of row standardizing would consider the spatial parameters to be in $(-1, 1)$. Since neither the Lee and Liu 2010 nor the Kelejian and Prucha 2008 parameter space concept cover or are useful to deduce the "practitioners" approach of row standardizing, it lacks a mathematically foundation. *Third*, this paper shows that in many applied cases the Lee and Liu 2010 parameter space is too restrictive.

These three problems of current parameter space concepts are the main motivation to define the spatial parameter space indirectly via desired mathematical properties. The properties are chosen so that the new spatial parameter space can be applied to draw inferences for example with mean distance estimators, like Generalized Method of Moments or Maximum Likelihood.

Due to this new parameter space definition the paper provides examples where accounting for the inner structure of the spatial weight matrix results in spatial parameter spaces that the literature would have treated as nonstationary. Additionally it is possible to show that the $(-1, 1)$ parameter space for a spatial lag after applying the "practitioners" approach of row standardizing, will result in almost all applied cases stationary DPGs.

The outline of the paper is the following: The next section briefly describes some spatial econometric DGPs, introduces the spatial parameter spaces of Kelejian and Prucha 2008 and Lee and Liu 2010 and then shows some fundamental problems of them. The problems of Section 2 motivate a different definition for spatial parameter spaces which is presented in Section 3. Section 4 uses the definitions to show the power of the new parameter space concept. The last section briefly concludes and summarizes this work. The appendix provides some useful theorems, lemmas and proofs like the Neumannseries and Gerschgorintheorem.

[3] The second practitioners approach, namely maximum row standardizing is covered by the parameterspace defnition of Lee and Liu 2010.

2 Important spatial DGPs and problems of the Kelejian & Prucha and the Lee & Liu parameter space

This section provides two general spatial econometric data generating processes, introduces the spatial parameter spaces of Kelejian and Prucha 2008 and Lee and Liu 2010 and describes their problems. For further details regarding the different DGPs see the references.

Notation: Let $\mathbf{Y}_n, \mathbf{S}_n \in \mathbb{R}^{n\times 1}$. One can write $\mathbf{Y}_n$ also as $\mathbf{Y} = (y_1, y_2, ..., y_n)'$. $|\chi|$ denotes the determinant of the matrix χ or the absolute value of the scalar χ, while $\|\|_\infty$ denotes the maximum absolute row sum and $\|\|_1$ the maximum absolute column sum of a matrix[4]. Due to notational convenience, not all indices will always be written. It should be clear from the context. $Sp(\mathbf{W}_n)$ denotes the set containing the eigenvalues of the matrix $\mathbf{W}_n$.

2.1 The Manski Model and the problems of Kelejian & Prucha parameter space

One general formulation for a spatial econometric DGP is represented by the Manski model (1).

$$\begin{gathered}\mathbf{Y}_n = \rho\mathbf{W}_n\mathbf{Y}_n + \mathbf{X}_n\beta + \mathbf{W}_n\mathbf{X}_n\theta + \mathbf{u}_n \\ \text{where } \mathbf{u}_n{=}\lambda\mathbf{W}_n\mathbf{u}_n{+}\boldsymbol{\epsilon}_n,\ \epsilon_i \sim i.i.d(0,\sigma^2)\end{gathered} \tag{1}$$

In (1) $\mathbf{X}_n$ represents the matrix of explanatory variables, ρ, β, θ and λ are the parameters to be estimated and the ϵ_i are independently and identically distributed with zero mean and finite variance σ^2. $\mathbf{W}_n$ represents the n by n spatial weights matrix of known constants. The diagonal entries of $\mathbf{W}_n$ are assumed to be zero[5]. The Manski model incorporates various representations of spatial DGPs like the Spatial Autoregressive Model, the Spatial Error Model and the Spatial Durbin Model[6]. The DGP stated in (1) can be solved for $\mathbf{Y}_n = (\mathbf{I}_n - \rho\mathbf{W}_n)^{-1}(\mathbf{X}_n\beta + \mathbf{W}_n\mathbf{X}_n\theta + (\mathbf{I}_n - \lambda\mathbf{W}_n)^{-1}\boldsymbol{\epsilon})$ if $(\mathbf{I}_n - \rho\mathbf{W}_n)^{-1}$ and $(\mathbf{I}_n - \lambda\mathbf{W}_n)^{-1}$ exist.

The parameter space proposed by Kelejian and Prucha 2008 for the Manski model stated in (1) can be sketched in the following manner: They argue correctly that if $\mathbf{W}_n$ is not normalized $(\mathbf{I}_n - \rho\mathbf{W}_n)^{-1}$ might not be defined for some values of $\rho \in (-1, 1)$. Therefore they

[4] Note that these matrix norms satisfy the following useful inequality: $\|\mathbf{A}_n\mathbf{B}_n\|_\vartheta \leq \|\mathbf{A}_n\|_\vartheta \|\mathbf{B}_n\|_\vartheta$ where $\vartheta \in \{1, \infty\}$ and $\mathbf{A}_n$ and $\mathbf{B}_n$ are n by n matrices. For more detais see Bronstein et al 2000 page 268.

[5] Allthough it is possible to derive paramerter spaces for W_n matrices where the diagonal elements are not zero, it is not common in applications...

[6] For more details to the assumptions and properties of the data generating process stated in (1), see Elhorst [1]

suggest that the parameter space of the spatial parameter should be $(-1/\tau_n, 1/\tau_n)$, where $\tau_n = \max_{1\leq i\leq n}\{|\nu_i| \text{ where } \nu_i \in Sp(\mathbf{W}_n)\}$. Since evaluating the eigenvalues of $\mathbf{W}_n$ can be numerically difficult, they suggest to use the Gershgorin theorem to get an upper bound for τ_n, what they call $\tau_n^* = \min\left\{\max_{1\leq i\leq n}\left\{\sum_{j=1}^{n}|w_{ij,n}|\right\}, \max_{1\leq i\leq n}\left\{\sum_{j=1}^{n}|w_{ji,n}|\right\}\right\}$ where $|\tau_n| \leq |\tau_n^*|$. As a result they recommend to use $(-1/\tau_n^*, 1/\tau_n^*)$ as a parameter space for the spatial parameter[7].

Although there are $\mathbf{W}_n$-matrices like the one forward one behind pattern[8], where guaranteeing that $\vee n \in \mathbb{N}\cup\{\infty\} : 1/\rho \notin Sp(\mathbf{W}_n)$ results in a well defined parameter space, this is generally not the case. Consider for example $\mathbf{W}_n = \overline{\mathbf{W}}_n$ where the typical element $\overline{w}_{ij}$ is defined by (2)

$$\overline{w}_{i,j} = \begin{cases} 1 \text{ if } i = j+1 \text{ and } j \in \{1,2,...,n-1\} \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

One could interpret the lag $\overline{\mathbf{W}}_n\mathbf{Y}_n$ in a model like $\mathbf{Y}_n = \rho\overline{\mathbf{W}}_n\mathbf{Y}_n+\boldsymbol{\varepsilon}_n$ as the time lag of a first order autoregressive process form the time series literature. Obviously the (admissible) parameter space for this "time" lag is $(-1,1)$ given the DPG has to be stationary. As the appendix shows $Sp\left(\overline{\mathbf{W}}_n\right) = \{0\}$. Therefore, the parameter space in Kelejian and Prucha 2008 would be $(-1/\tau_n, 1/\tau_n) = (-\infty,\infty) = \mathbb{R}$ and hence can result in nonstationary DGPs. This is not surprising, since there is a difference between the solvability and the boundedness of a spatial DGP. For the $\overline{\mathbf{W}}_n$ stated in (2), one can write down an analytical solution for the inverse:[9] $\vee n \in \mathbb{N} : \left(\mathbf{I}_n - \rho\overline{\mathbf{W}}_n\right)^{-1} = \sum_{k=0}^{n}\rho^k\overline{\mathbf{W}}_n^k$. Note that the series $\sum_{k=0}^{n}\rho^k\overline{\mathbf{W}}_n^k$ converges for every $n\in\mathbb{N}$. This example shows if the DGP stated in (1) is not solvable for a $n\in\mathbb{N}$, it is also not stable, but if the DGP is solvable it does not mean that it is stable as $n\to\infty$. Therefore, one has to use additional assumptions for the parameter space like the boundedness[10] of $\left(\mathbf{I}_n - \rho\overline{\mathbf{W}}_n\right)^{-1}$ in absolute row and column sums as $n\to\infty$. Since $\sum_{k=0}^{\infty}\rho^k\overline{\mathbf{W}}_n^k$ converges only if $|\rho|<1$, the boundedness condition constrains the spatial parameter space in a useful manner.

[7] Please note that this parameter space is not the same Lee and Liu 2010 proposed.

[8] see Koch 2010

[9] Note that in this case $\left(\overline{\mathbf{W}}_n\right)^{n+1} = \mathbf{0}_{n\times n}$, where $\mathbf{0}_{n\times n}$ denotes a matrix of size n by n only containing zeros.

[10] This is a standard assumption for drawing inferences on spatial econometric models, see for example ....

2.2 High order spatial models and the problems of Lee & Liu parameter space

Another general spatial econometric DGP[11] is represented by the high order spatial autoregressive models with autoregressive disturbances[12].

$$\begin{gathered}\mathbf{Y}_n=\sum_{j=1}^{p}\rho_j\mathbf{W}_{jn}\mathbf{Y}_n+\mathbf{X}_n\beta+\mathbf{u}_n \\ \text{where } \mathbf{u}_n=\sum_{j=1}^{q}\lambda_k\mathbf{M}_{kn}\mathbf{u}_n+\boldsymbol{\epsilon}_n,\ \epsilon_i\sim i.i.d(0,\sigma^2)\end{gathered} \tag{3}$$

In (3) $\rho_1,...,\rho_p$ and $\lambda_1,...,\lambda_q$ represent the different spatial lag parameters, $\mathbf{W}_{1n},...,\mathbf{W}_{pn}$ and $\mathbf{M}_{1n},...,\mathbf{M}_{qn}$ are n by n dimensional spatial weights matrices. Like in (1) it is assumed that the diagonal elements of $\mathbf{W}_{jn}$ and $\mathbf{M}_{kn}$ are set to zero. The DGP stated in (3) can be solved for $\mathbf{Y}_n=(\mathbf{I}_n-\sum_{j=1}^{p}\rho_j\mathbf{W}_{jn})^{-1}\ (\mathbf{X}_n\beta+\ \ (\mathbf{I}_n-\sum_{k=1}^{q}\lambda_k\mathbf{M}_{kn})^{-1}\boldsymbol{\epsilon})$ if $\mathbf{I}_n-\sum_{j=1}^{p}\rho_j\mathbf{W}_{jn}$ and $\mathbf{I}_n-\sum_{k=1}^{q}\lambda_k\mathbf{M}_{kn}$ are invertible.

While Kelejian and Prucha 2008 motivate their parameter space via the eigenvalues of $\mathbf{W}_n$ Lee and Liu 2010 on the other hand motivate the spatial parameter space for the DGP stated in (3) via the following reasoning: They argue correctly if $\sum\limits_{j=1}^{p}\left|\rho_j\right|<1/\max_{j=1...p}\{\|\mathbf{W}_{jn}\|_1,\|\mathbf{W}_{jn}\|_\infty\}$ and $\sum\limits_{k=1}^{q}|\lambda_k|<1/\max_{k=1...q}\{\|\mathbf{M}_{kn}\|_1,\|\mathbf{M}_{kn}\|_\infty\}$ that the inverses of $\mathbf{I}_n-\sum\limits_{j=1}^{p}\rho_j\mathbf{W}_{jn}$ and $\mathbf{I}_n-\sum\limits_{j=1}^{q}\lambda_k\mathbf{M}_{kn}$ exist due to the existence of the Neumannseries and both are bounded in row and column sums in absolute value. This condition represents their spatial parameter space. Although it can be applied on neighboring patterns like (2) the parameter space can be restrictive if $p=1,q=0$ and $\mathbf{W}_n$ is row normalized.

As an example to show the constraints of the Lee and Liu 2010 spatial parameter space, the spatial one forward one behind pattern represented by $\overline{\mathbf{W}}_n$ is used where the typical element $\overline{w}_{i,j}$ is defined by (4)

$$\overline{w}_{i,j}=\begin{cases}\overline{w}_{j,i}=1 \text{ if } j=i+1 \text{ and } i\in\{1,2,...,n-1\}\\ 0 \text{ otherwise}\end{cases} \tag{4}$$

[11] [see Lee and Liu 2010 for model details]

[12] Lee and Liu [2010] have suggested to use either a ML or a B-IV estimator for this DGP. Badinger and XXX [] have suggested to use an IV- estimator.

Row-normalizing $\mathbf{W}_n$ yields $\widetilde{\mathbf{W}}_n$. Obviously $\left\|\widetilde{\mathbf{W}}_n\right\|_\infty = 1$ holds, but $\left\|\widetilde{\mathbf{W}}_n\right\|_1 = 1.5$. Therefore, it is not clear, whether $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_1 = \left\|\sum_{k=0}^{\infty} \rho_n^k \widetilde{\mathbf{W}}_n^k\right\|_1$ converges, because $\left\|\widetilde{\mathbf{W}}_n\right\|_1 > 1$. Hence, the parameter space stated in Lee and Liu 2010 for $\widetilde{\mathbf{W}}_n$ would be $\rho \in \left(-\frac{1}{1.5}, \frac{1}{1.5}\right) = \left(-\frac{2}{3}, \frac{2}{3}\right)$ in order to ensure convergence of $\left\|\sum_{k=0}^{\infty} \rho_n^k \widetilde{\mathbf{W}}_n^k\right\|_1$.

The previous reasoning does not imply that the spatial parameter space of (4) has to be $\left(-\frac{2}{3}, \frac{2}{3}\right)$. As Theorem 1 in section 4.1 in shows that for row standardized $\mathbf{W}_n$-matrices like $\widetilde{\mathbf{W}}_n$ given in (4) the parameter space is still $(-1, 1)$. Therefore, the Lee and Liu 2010 parameter space results into too restrictive parameter spaces for all $\mathbf{W}_n$ where a row standardization was applied and the row sums of $\mathbf{W}_n$ before the normalization were different and the original weight matrix was symmetric.

3 A formal Parameter space definition

In order to derive estimator properties for DGPs like (1) or (3), stationarity is required. The stationarity assumption is reflected via boundedness conditions of the DGP. Hence, this paper proposes to use these conditions to construct the spatial parameter space. Consequently the proposed spatial parameter space has to satisfy the following properties 1, 2, and 3:

Definition 1 *Let* $\rho_i \in \Theta_i \subset \mathbb{R}$, $\mathbf{\Psi}_{n,p} = \sum\limits_{j=1}^{p} \rho_j \mathbf{W}_{jn}$ *if* $p > 1$, $\mathbf{\Psi}_{n,1} = \rho\mathbf{W}_n$ *and the diagonal entries of* $\mathbf{W}_n$ *be zero.* Θ_i *where* $i \in \{1, ..., p\}$ *is a well defined spatial parameter space, if the following properties are met:*

1. $\forall \rho_i \in \Theta_i, n \in \mathbb{N} \cup \{\infty\} : |\mathbf{I}_n - \mathbf{\Psi}_{n,p}| \neq 0$
2. $\forall \rho_i \in \Theta_i, n \in \mathbb{N} \cup \{\infty\} : \left\|(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1}\right\|_1 < \infty \wedge \left\|(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1}\right\|_\infty < \infty \wedge \left\|\mathbf{\Psi}_{n,p}\right\|_\infty < \infty \wedge \left\|\mathbf{\Psi}_{n,p}\right\|_1 < \infty$
3. $\Theta_i = \bigcup\limits_{j \in \mathbb{N}} A_j$, A_j *is an interval of* $\mathbb{R}$.

The first property simply states that for every sample size, even if it approaches infinity, there must always exist the inverse of $\mathbf{I}_n - \rho\mathbf{W}_n$ or $\mathbf{I}_n - \sum_{j=1}^{p} \rho_j \mathbf{W}_{jn}$. If one would only use the the first condition in order to find parameter spaces, one would use the Kelejian Prucha 2008 parameter space. The first condition simply ensures that the spatial DPG exists for every $n \in \mathbb{N} \cup \{\infty\}$.

The second property ensures the boundedness of the inverse in absolute row and column sums. In an econometric this property ensures finite moments of $\mathbf{Y}_n$. This property guarantees for $\mathbf{W}_n$ as defined in equation (2) that the parameter space is only $(-1, 1)$. The properties $\|\mathbf{\Psi}_{n,p}\|_\infty < \infty \wedge \|\mathbf{\Psi}_{n,p}\|_1 < \infty$ imply that the spatial spillover has to be bounded as well.

The third property reflects the idea that only intervals are used as a parameter space and not a countable set of points. This is for example necessary if mean value theorems for deriving estimator properties are used.

Properties 1 and 2 clearly show the difference between solvability and stationarity of a spatial DGP. Although there exist examples, like the spatial one forward one behind pattern, where property 1 is sufficient for property 2, this is not generally the case.

Equipped with the properties 1- 3 the parameter space proposed by Lee and Liu 2010 becomes clearer. If there is no additional knowledge about $\mathbf{W}_{jn}$ and an explicit spatial parameter space for ρ_i is desired one could use the following reasoning: If any matrix norm of $\mathbf{\Psi}_{n,p}$ is smaller than one the Neumann series can be applied to find the inverse of $\mathbf{I}_n - \mathbf{\Psi}_{n,p}$. Additionally Lemma 1 shows for $\varrho \in \{1, \infty\}$ if $\|\mathbf{\Psi}_{n,p}\|_\varrho < 1$ that $\left\|(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1}\right\|_\varrho \leq \frac{1}{1-\|\mathbf{\Psi}_{n,p}\|_\varrho} < \infty$. This equivalent to the spatial parameter space suggested by Lee and Liu 2010.

The next section shows, how these properties 1-3 can help to construct spatial parameter spaces that are larger than the ones currently considered by the literature. Additionally the next section shows which conditions are necessary in order for the "practitioners" approach of row normalizing to fulfill the properties 1-3.

4 Power of the new parameter space concept

This section provides three applications of the new parameter space concept in order to show its power. Since the concept is defined indirectly via mathematical properties it is possible to use the inner structure of the spatial lag(s) to derive the corresponding spatial parameter spaces. The examples of this section suggest that the more inner structure of the spatial lag(s) is present the more precise and in these examples larger the spatial parameter space gets. *First* subsection 4.1 delivers the mathematical foundation for the "practitioners" approach of row standardizing. The key assumption for this proof is that the $\mathbf{W}_n$ was symmetric before the row standardization took place. A *second* application is a group interaction model where the within and between group interaction is modelled with partitioned $\mathbf{W}$- matrices. This particular inner structure of the spatial lags results in a significantly larger spatial parameter space than the one proposed by Lee and Liu 2010. Subsection 4.3 provides the *third* application, where it turns out that in some panel settings it is possible to use almost the whole real line as the spatial parameter space. Additionally this section suggests one possible interpretation for the inevitable cross section sample size dependence of the parameter space. In the context of repeated sampling it can be seen for example as a consequence of the geographic scale.

4.1 Row standardizing and Stability

This subsection shows that the "practitioners" approach of row standardizing yields stationary parameter spaces. Let $\mathbf{W}_n$ be the spatial weight matrix with the typical element $w_{i,j}$ and $w_{i,i} = 0$. The row standardization for $\mathbf{W}_n$ is represented by the diagonal matrix $\mathbf{\Lambda}_n$ with the typical element[13] $\lambda_{i,i} = 1/\sum_{j=1}^{n} |w_{i,j}|$. Hence the normalized weight matrix is given by $\widetilde{\mathbf{W}}_n = \mathbf{\Lambda}_n \mathbf{W}_n$. The aim is to find weather $\rho \in (-1, 1)$ fulfills the parameter space properties 1-3. Property 3 is obviously fulfilled. In order to find the inverse of $\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n$ a Numannseries can be applied $(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n)^{-1} = \sum_{k=0}^{\infty} \rho^k \widetilde{\mathbf{W}}_n^k$ and hence, property (1) is fulfilled. To show property (2) via Theorem 1 additional assumptions are necessary: $\mathbf{W}_n$ has to be symmetric and $\|\mathbf{W}_n\|_\infty < \kappa \in \mathbb{R}$.

The first assumption of a symmetric $\mathbf{W}_n$ is generally fulfilled if for example, $\mathbf{W}_n$ represents a spatial neighboring structure. If observation A is neighbor of observation B, the reverse must also be true. The second assumption requests that $\mathbf{W}_n$ has bounded absolute row sums. In the context of a neighboring structure that is equivalent to limiting the number of neighbors for each observation to a finite constant.

Theorem 2 *Let $\mathbf{W}_n$ be a symmetric weigh matrix with finite weights $w_{i,j} \in \mathbb{R}$. The dependence structure is limited such that $\|\mathbf{W}_n\|_\infty < \kappa \in \mathbb{R}$. Let $\widetilde{\mathbf{W}}_n$ be the row standardized version of $\mathbf{W}_n$. If $|\rho| < 1$ it follows that:* $\left\| \left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1} \right\|_\infty \wedge \left\| \left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1} \right\|_1 \wedge \left\|\widetilde{\mathbf{W}}_n\right\|_\infty \wedge \left\|\widetilde{\mathbf{W}}_n\right\|_1 \leq c \in \mathbb{R}$

Proof. 4 Properties have to be shown:

[13] It is assumed that $\sum_{j=1}^{n} |w_{i,j}| > 0$. If $w_{i,j}$ represents a neigboring structur, $\sum_{j=1}^{n} |w_{i,j}| > 0$ is fullfilled if each observation has at least one neigbor.

(1)$\left\|\widetilde{\mathbf{W}}_n\right\|_\infty < \infty$: Due to the construction of $\mathbf{\Lambda}_n$: $\left\|\widetilde{\mathbf{W}}_n\right\|_\infty = \|\mathbf{\Lambda}_n\mathbf{W}_n\|_\infty$ $= 1 < \infty$.

(2)$\left\|\widetilde{\mathbf{W}}_n\right\|_1 < \infty$: Observe that $\left\|\widetilde{\mathbf{W}}_n\right\|_1 = \|\mathbf{\Lambda}_n\mathbf{W}_n\|_1 = \|\mathbf{W}_n'\mathbf{\Lambda}_n\|_\infty$ holds. Since $\mathbf{W}_n$ is symmetric and the property of the sub- multiplicatively of matrix norms $\|\mathbf{W}_n'\mathbf{\Lambda}_n\|_\infty = \|\mathbf{W}_n\mathbf{\Lambda}_n\|_\infty \leq \|\mathbf{W}_n\|_\infty \|\mathbf{\Lambda}_n\|_\infty < \kappa\lambda_{\max} < \infty$ where $\lambda_{\max} = \max_{1\leqslant i \leqslant n}\left\{1/\sum_{j=1}^n |w_{i,j}|\right\}$.

(3) $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_\infty < \infty$: Due to Lemma YYY (applying the Neumannseries and using the subadditivity/submultiplicativity of matrix norms) it follows that: $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_\infty = \frac{1}{1-|\rho|} \leq c \in \mathbb{R}$.

(4) $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_1 < \infty$: First a Neumannseries and the subadditivity of matrix norms is used: $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_1 \leq \sum\limits_{k=0}^{\infty} |\rho|^k \left\|\left(\widetilde{\mathbf{W}}_n\right)^k\right\|_1$.

Note that if $k > 2$: $\left\|\widetilde{\mathbf{W}}_n^k\right\|_1 = \left\|\mathbf{\Lambda}_n\left(\prod\limits_{l=1}^{k-1}\mathbf{W}_n\mathbf{\Lambda}_n\right)\mathbf{W}_n\right\|_1 \leq \|\mathbf{\Lambda}_n\|_1 \left\|\prod\limits_{l=1}^{k-1}\mathbf{W}_n\mathbf{\Lambda}_n\right\|_1$ $\|\mathbf{W}_n\|_1 = \|\mathbf{\Lambda}_n\|_1 \left\|\prod\limits_{l=1}^{k-1}\mathbf{\Lambda}_n\mathbf{W}_n\right\|_\infty \|\mathbf{W}_n\|_1 = \lambda_{\max}\kappa$ holds. Hence $\left\|\left(\mathbf{I}_n - \rho\widetilde{\mathbf{W}}_n\right)^{-1}\right\|_1$ $\leq 1 + \sum\limits_{k=1}^{\infty} |\rho|^k \kappa\lambda_{\max} < \infty$ ■

4.2 Group- interactions and spatial econometric modelling

This subsection considers the following empirical problem: The DGP not only contains spatial autocorrelation but also has to account for different spatial interaction parameters, namely within and between groups. The DGP could for example model a housing market with two distinct geographical markets, namely an east-market and a west-market. These two groups (east and west) have the sample sizes n_1 (west- market) and n_2 (east- market). The overall sample size is denoted by $n = n_1 + n_2$. For simplicity it is assumed that the data is ordered by these particular groups. To account for the within and between group effects the following model specification (5) could be used[14].

$$\begin{gathered}\mathbf{Y}_n = \left(\rho_{11}\hat{\mathbf{W}}_{11} + \rho_{12}\hat{\mathbf{W}}_{12} + \rho_{21}\hat{\mathbf{W}}_{21} + \rho_{22}\hat{\mathbf{W}}_{22}\right)\mathbf{Y}_n + \mathbf{X}_n\widetilde{\beta} + \boldsymbol{\epsilon}_n \\ \text{where } \epsilon_i \sim i.i.d(0, \sigma^2)\end{gathered} \tag{5}$$

[14] Of course in order to find asymptotic properties for the spatial parameters it has to assumed that n_1 and n_2 $\to \infty$

Since the data is ordered, the spatial weight matrices have the following simple form: $\hat{\mathbf{W}}_{11} = \begin{pmatrix} \mathbf{W}_{n1,n1} & \mathbf{0}_{n1,n2} \\ \mathbf{0}_{n2,n1} & \mathbf{0}_{n2,n2} \end{pmatrix}$, $\hat{\mathbf{W}}_{12} = \begin{pmatrix} \mathbf{0}_{n1,n1} & \mathbf{W}_{n1,n2} \\ \mathbf{0}_{n2,n1} & \mathbf{0}_{n2,n2} \end{pmatrix}$, $\hat{\mathbf{W}}_{21} = \begin{pmatrix} \mathbf{0}_{n1,n1} & \mathbf{0}_{n1,n2} \\ \mathbf{W}_{n2,n1} & \mathbf{0}_{n2,n2} \end{pmatrix}$, $\hat{\mathbf{W}}_{22} = \begin{pmatrix} \mathbf{0}_{n1,n1} & \mathbf{0}_{n1,n2} \\ \mathbf{0}_{n2,n1} & \mathbf{W}_{n2,n2} \end{pmatrix}$. Note that $\hat{\mathbf{W}}_{11}$, $\hat{\mathbf{W}}_{12}$,$\hat{\mathbf{W}}_{21}$ and $\hat{\mathbf{W}}_{22}$ are partitioned.

The within group effects, given by the terms $\rho_{11}\hat{\mathbf{W}}_{11}\mathbf{Y}$ and $\rho_{22}\hat{\mathbf{W}}_{22}\mathbf{Y}$ would answer the question of how are the western- $(\rho_{11}\hat{\mathbf{W}}_{11}\mathbf{Y})$ and eastern- market $(\rho_{22}\hat{\mathbf{W}}_{22}\mathbf{Y})$ affected by themselves. On the other hand the between groups effects, given by the terms $\rho_{12}\hat{\mathbf{W}}_{12}\mathbf{Y}$ and $\rho_{21}\hat{\mathbf{W}}_{21}\mathbf{Y}$ would answer the question of how is the western market influenced by the eastern $(\rho_{12}\hat{\mathbf{W}}_{12}\mathbf{Y})$ market and vice versa $(\rho_{21}\hat{\mathbf{W}}_{21}\mathbf{Y})$.

To find the spatial parameter space for (5) let $\mathbf{\Psi}_{n,4}$ be defined by $\mathbf{\Psi}_{n,4} := \rho_{11}\hat{\mathbf{W}}_{11} + \rho_{12}\hat{\mathbf{W}}_{12} + \rho_{21}\hat{\mathbf{W}}_{21} + \rho_{22}\hat{\mathbf{W}}_{22}$. Additionally it is assumed[15] that $\max\{||\hat{\mathbf{W}}_{i,j}||_{\varrho}$ where $\varrho \in \{1,\infty\}$ and $i,j \in \{1,2\}\} \leq 1$. Although not necessary for the following proof, in most applied cases $\mathbf{W}_{n1,n1}$ and $\mathbf{W}_{n2,n2}$ will be symmetric and $\mathbf{W}_{n1,n2} = (\mathbf{W}_{n2,n1})'$.

The following paragraph will sketch the search and proof for the parameter space, the actual proof is given in the appendix: The first step is to use the equation system $\mathbf{y}_1 = \rho_{11}\mathbf{W}_{n1,n1}\mathbf{y}_1 + \rho_{12}\mathbf{W}_{n1,n2}\mathbf{y}_2 + \mathbf{s}_1$ (West) and $\mathbf{y}_2 = \rho_{21}\mathbf{W}_{n2,n1}\mathbf{y}_1 + \rho_{22}\mathbf{W}_{n2,n2}\mathbf{y}_2 + \mathbf{s}_2$ (East) and solve for $\mathbf{y}_1$ and $\mathbf{y}_2$. One condition for solving these equations is that $\mathbf{A}_1 := (\mathbf{I}_{n1} - \rho_{11}\mathbf{W}_{n1,n1})^{-1}$ and $\mathbf{A}_2 := (\mathbf{I}_{n2} - \rho_{22}\mathbf{W}_{n2,n2})^{-1}$ are defined. Therefore, two conditions for the parameter space are: $|\rho_{11}| < 1$ and $|\rho_{22}| < 1$. The second step is to find the restrictions for ρ_{21} and ρ_{12}. Since it is assumed that $|\rho_{11}| < 1$ and $|\rho_{22}| < 1$ one can solve the two equations $\mathbf{y}_1 = (\mathbf{I}_{n1} - \rho_{11}\mathbf{W}_{n1,n1})^{-1}(\rho_{12}\mathbf{W}_{n1,n2}\mathbf{y}_2 + \mathbf{s}_1)$ and $\mathbf{y}_2 = (\mathbf{I}_{n2} - \rho_{22}\mathbf{W}_{n2,n2})^{-1}(\rho_{21}\mathbf{W}_{n2,n1}\mathbf{y}_1 + \mathbf{s}_2)$. Inserting $\mathbf{y}_2$ into $\mathbf{y}_1$ yields: $\mathbf{y}_1 = \mathbf{A}_1(\rho_{12}\mathbf{W}_{n1,n2}\mathbf{A}_2(\rho_{21}\mathbf{W}_{n2,n1}\mathbf{y}_1 + \mathbf{s}_2) + \mathbf{s}_1)$. It is possible to solve this equation for $\mathbf{y}_1$ if $\|\mathbf{A}_1\rho_{12}\mathbf{W}_{n1,n2}\mathbf{A}_2\rho_{21}\mathbf{W}_{n2,n1}\|_1 < 1$. This inequality is satisfied if $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$. Hence it is shown that the following two conditions fulfill the parameter space properties 1-3 for the DGP given in (5).

1. $|\rho_{11}| < 1$ and $|\rho_{22}| < 1$

2. $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$

[15] The appendix shows that if $\mathbf{W}_{n1,n1}$, $\mathbf{W}_{n2,n2}$, $\mathbf{W}_{n2,n1}$ and $\mathbf{W}_{n1,n2}$ are row standardized, $\mathbf{W}_{n1,n1}$ and $\mathbf{W}_{n2,n2}$ can be written as $\mathbf{W}_{n1,n1} = \Lambda_{n1}\overline{\mathbf{W}}_{n1,n1}$ and $\mathbf{W}_{n2,n2} = \Lambda_{n2}\overline{\mathbf{W}}_{n2,n2}$ where Λ represents the row- standardizing and both $\overline{\mathbf{W}}$ are symmetric and $\|\overline{\mathbf{W}}\|_\infty, \|\mathbf{W}_{n1,n2}\|_\infty, \|\mathbf{W}_{n2,n1}\|_\infty < \infty$ hold the same spatial parameter space can be applied for the DGP given in (6)

This parameter space is significantly larger than the one proposed

by Lee and Liu 2010. For example if $|\rho_{11}| = |\rho_{22}| = 1/3$ the "" area in figure (1) represents the Lee and Liu parameter space and "" represents the parameter space due to inequality $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$.

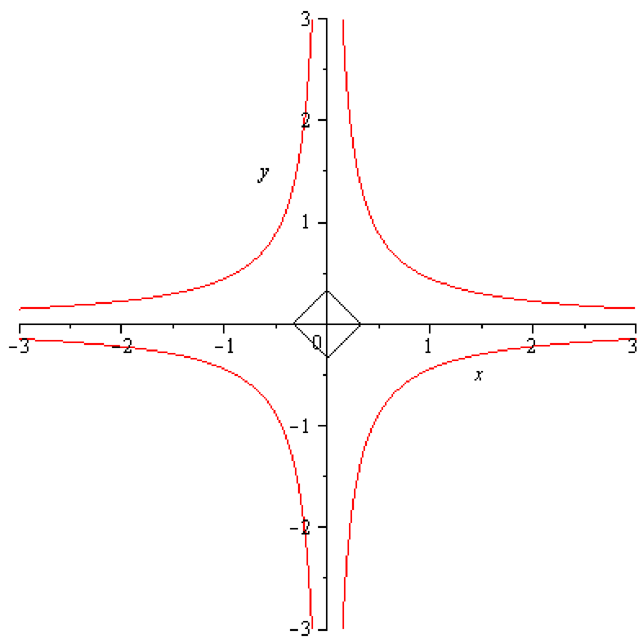

figure 1:....

Figure (1) shows quite dramatically that one should take the inner structure of the $\mathbf{W}_n$- matrices into account in order to find parameter spaces. This is to some extend similar to the "practitioners" approach of row standardizing. The more knowledge is present about the inner structure of the spatial lags, like in this example the weight matrices are partitioned, the more precise and in this particular example larger is the resulting spatial parameter space gets. This can also be seen in the next application of the new parameter space definition.

4.3 Geographic scale and the spatial parameter space

This subsection explores panels where the number of observation units (n) are fixed. Since n is fixed T the number of time periods has to go to infinity in order to derive the asymptotic properties of possible estimators. Let N denote the overall sample size, namely $N = nT$. The vector $\mathbf{Y}_N$ has the elements $y_{1,1}, y_{2,1}, ..., y_{n,1}, ..., y_{n,T}$. A possible spatial DGP reflecting these properties is given by (6)

$$\mathbf{Y}_N = \rho\left(\mathbf{I}_T \otimes \mathbf{W}_n\right)\mathbf{Y}_N + \mathbf{X}_N\beta + \boldsymbol{\epsilon}_N \text{ where } \epsilon_i \sim i.i.d(0, \sigma^2) \tag{6}$$

where $\otimes$ denotes the Kronecker product. Finding a spatial parameter space for (6) can be achieved by the following reasoning: Obviously it has to be ensured that the inverse of $\mathbf{I}_N - \rho\left(\mathbf{I}_T \otimes \mathbf{W}_n\right)$ exists. Since $Sp\left(\mathbf{I}_T \otimes \mathbf{W}_n\right) = Sp\left(\mathbf{W}_n\right)$ the condition $1/\rho \notin Sp(\mathbf{W}_n)$ follows directly. Note that the number of elements in $Sp(\mathbf{W}_n)$ is always smaller or equal n. Since n is fixed $\|\mathbf{W}_n\|_\infty < \infty$, $\|\mathbf{W}_n\|_1 < \infty$ and $\left\|(\mathbf{I}_n - \rho\mathbf{W}_n)^{-1}\right\|_{1,\infty} < \infty$ have to hold for $1/\rho \notin Sp(\mathbf{W}_n)$. Remark that $\left(\mathbf{I}_N - \rho\left(\mathbf{I}_T \otimes \mathbf{W}_n\right)\right)^{-1} = \mathbf{I}_T \otimes \left(\mathbf{I}_n - \rho\mathbf{W}_n\right)^{-1}$ and therefore $\left\|\left(\mathbf{I}_N - \rho\left(\mathbf{I}_T \otimes \mathbf{W}_n\right)\right)^{-1}\right\|_{1,\infty} < \infty$ for all $T \in \mathbb{N} \cup \{\infty\}$ as long $1/\rho \notin Sp(\mathbf{W}_n)$. Thus the spatial parameter space for (6) is given by $\mathbb{R}\backslash\{1/\widetilde{\tau}_n$ where $\widetilde{\tau}_n \in Sp(\mathbf{W}_n)\}$.

The previous paragraph showed above all two noteworthy characteristics of the spatial parameter space for the DGP given in (6): First, $\mathbb{R}\backslash\{1/\widetilde{\tau}_n$ where $\widetilde{\tau}_n \in Sp(\mathbf{W}_n)\}$ is tremendously larger than the parameter spaces considered by the literature so far. Second, the spatial parameter space is a function of the by assumption fixed n and thus can offer an interesting interpretation: The implicit n dependence of the DGP dynamics can for example be seen as a consequence of different geographic scales. To illustrate this point let $n = 27$ represent the countries in the European Union and $y_{i,t}$ their corresponding GDP- growth rate. If n (and hence $\mathbf{W}_n$) is now changed to $\widetilde{n} = 271$ ($\widetilde{n}$ now represents the NUTS-2 regions of the European Union) this simply would reflect a change in the geographical scale. Since a change in the geographical scale is often associated with a change in the model dynamics a parameter space as function dependent on n seems plausible and would simply reflect some influences of the geographical scale on the DGP.

It is important that the previous reasoning is not suggesting to use $\mathbb{R}\backslash\{1/\widetilde{\tau}_n$ where $\widetilde{\tau}_n \in Sp(\mathbf{W}_n)\}$ for every spatial panel data DGP. Whether this makes sense, depends on what the DGP should describe. Consider for example house prices in a real estate market. Economic theory would suggest that the market prices are independent of the market size n. As a result the DPG- dynamics should be independent of n as well and consequently the parameter space must be independent of n too, like for example the classical $(-1, 1)$- spatial parameter space, after normalizing the weigth matrix.

5 Conclusion and Summary

Unlike the time series literature, there has been not much effort in the spatial econometric literature to substantially examine the parameter spaces for spatial econometric models. This paper raises three important issues concerning spatial parameter spaces:

First, current parameter space concepts and practical approaches are inadequate. This point is supported by the three following observations:

- Since the Kelejian and Prucha 2008 parameter space is only considering the eigenvalues of the spatial weight matrix, it is only concerned about the existence of the DGP and not its stationarity. Hence it can result in nonstationary DGPs if for example the spatial weight matrix mimic a process from the time series literature.

- The Lee and Liu 2010 parameter space can result in too small parameter spaces, especially if it is confronted with row standardized weight matrices. Although it will always result in stationary DGPs it has to be seen as too restrictive.

- Since neither the Kelejian and Prucha 2008 nor the Lee and Liu 2010 parameter space can be used as a mathematical foundation for the "practitioners" approach of row standardizing, this approach lacks a theoretical basis.

Second, a useful spatial parameter space can be defined indirectly via desired mathematical properties. These properties are showing clearly the difference between the necessary conditions for the existence of the DGP and its stationarity. Additionally it shows that the Lee and Liu 2010 parameter space can be seen as a special case of this new parameter space definition.

Third, the power of the new parameter space concept lies in its ability to account for the inner structure of the spatial lag(s). Hence it is possible to derive more precise and in some cases larger spatial parameter spaces. This can be verified with the help of three practical examples:

- Section 4.1 shows that the "practitioners" approach of row standardizing under the practical assumption that if the weight matrix before row normalizing was symmetric, the approach always yields stationary DGPs. Hence, it is possible to give the "practitioners" approach of row standardizing a mathematical foundation.

- Section 4.2 handles models with different spatial interactions reflecting the assumed group structure. It is possible to find substantially larger spatial parameter spaces than the ones previously considered by the literature.

- Finally, the example of spatial panels with fixed n raises two interesting issues: First it is showing that under certain assumptions about the DGP, it is possible to use almost the whole real line as the spatial parameter space. Second, it suggested that the implicit n dependence of the DGP dynamics can for example be seen as a consequence of different geographic scales.

These results highlight the importance of the spatial parameter space. Therefore, applied reseachers should be encouraged to deal with their parameter space in more detail since it could be larger and reveal some dynamics of the spatial DGP.

A Appendix

Lemma 3 *Let* $\overline{\mathbf{W}}_n$ *a* n *by* n *matrix where the typical element* $\overline{w}_{ij,n}$ *is defined by (2). It follows that:* $Sp(\overline{\mathbf{W}}_n) = \{0\}$.

Proof. We proof this theorem by using the Jordan normal form: $\overline{\mathbf{W}}_n = F_n \mathbf{\Gamma}_n F_n^{-1}$. If the typical element $f_{i,j}$ of F_n and $\gamma_{i,j}$ of $\mathbf{\Gamma}_n$ are defined by (7) and (8) it follows due to the Jordan normal form that the diagonal elements of $\mathbf{\Gamma}_n$ are the eigenvalues of $\overline{\mathbf{W}}_n$.

$$f_{i,j} = \begin{cases} 1 \text{ if } i = n+1-j \text{ and } j \in \{1,2,...,n\} \\ 0 \text{ otherwise} \end{cases} \tag{7}$$

$$\gamma_{i,j} = \begin{cases} 1 \text{ if } j = i+1 \text{ and } i \in \{1,2,...,n-1\} \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

Note that $F_n = F_n^{-1}$ and $\mathbf{\Gamma}_n$ is a typical Jordan form. It can easily be seen that $\overline{\mathbf{W}}_n = F_n \mathbf{\Gamma}_n F_n^{-1} = F_n \mathbf{\Gamma}_n F_n$ holds and hence $Sp(\overline{\mathbf{W}}_n) = \{0\}$. ■

Proof. for the group model -Part A: If we have a model like (5) the following parameter space fulfills all the 3 parameter space properties: $|\rho_{11}| < 1$, $|\rho_{22}| < 1$ and $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$. Let $\mathbf{\Psi}_{n,4} = \rho_{11}\hat{\mathbf{W}}_{11} + \rho_{12}\hat{\mathbf{W}}_{12} + \rho_{21}\hat{\mathbf{W}}_{21} + \rho_{22}\hat{\mathbf{W}}_{22}$. Due to $|\rho_{11}| < 1$ $|\rho_{22}| < 1$ and $\max\{||\hat{\mathbf{W}}_{i,j}||_\varrho$ where $\varrho \in \{1,\infty\}$ and $i,j \in \{1,2\}\} \leq 1$ the inverse of $\mathbf{I}_{n1} - \rho_{11}\mathbf{W}_{n1,n1}$ and $\mathbf{I}_{n2} - \rho_{22}\mathbf{W}_{n2,n2}$ exist. We use the equation system $\mathbf{y}_1 = \rho_{11}\mathbf{W}_{n1,n1}\mathbf{y}_1 + \rho_{12}\mathbf{W}_{n1,n2}\mathbf{y}_2 + \mathbf{s}_1$ and $\mathbf{y}_2 = \rho_{21}\mathbf{W}_{n2,n1}\mathbf{y}_1 + \rho_{22}\mathbf{W}_{n2,n2}\mathbf{y}_2 + \mathbf{s}_2$. We can solve these equations and get:

$\mathbf{y}_1 = (\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1})^{-1}\mathbf{A}_1(\rho_{12}\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{s}_2 + \mathbf{s}_1)$
$\mathbf{y}_2 = (\mathbf{I}_{n2} - \rho_{12}\rho_{21}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\mathbf{W}_{n1,n2})^{-1}\mathbf{A}_2(\rho_{21}\mathbf{W}_{n2,n1}\mathbf{A}_1\mathbf{s}_1 + \mathbf{s}_2)$

where $\mathbf{A}_1 = (\mathbf{I}_{n1} - \rho_{11}\mathbf{W}_{n1,n1})^{-1}$ and $\mathbf{A}_2 = (\mathbf{I}_{n2} - \rho_{22}\mathbf{W}_{n2,n2})^{-1}$. Therefore, we find $\mathbf{\Psi}_{n,4}^{-1} = \begin{pmatrix} \widetilde{\Psi}_{11} & \widetilde{\Psi}_{12} \\ \widetilde{\Psi}_{21} & \widetilde{\Psi}_{22} \end{pmatrix}$ where

$\widetilde{\Psi}_{11} = (\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1})^{-1}\mathbf{A}_1$,
$\widetilde{\Psi}_{12} = (\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1})^{-1}\mathbf{A}_1\rho_{12}\mathbf{W}_{n1,n2}\mathbf{A}_2$,
$\widetilde{\Psi}_{21} = (\mathbf{I}_{n2} - \rho_{12}\rho_{21}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\mathbf{W}_{n1,n2})^{-1}\mathbf{A}_2\rho_{21}\mathbf{W}_{n2,n1}\mathbf{A}_1$ and
$\widetilde{\Psi}_{22} = (\mathbf{I}_{n2} - \rho_{12}\rho_{21}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\mathbf{W}_{n1,n2})^{-1}\mathbf{A}_2$. If we use a Neumannseries we can show that $\mathbf{\Psi}_{n,4}^{-1}$ exists if $\|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n2,n1}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_1 \vee \|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n2,n1}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_\infty < 1$. $\|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n2,n1}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_1 \vee \|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_\infty < 1$ is true if $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$, since Lemma 6 shows $\|\mathbf{A}_1\|_{1,\infty} < \frac{1}{|1-\rho_{11}|}$and $\|\mathbf{A}_2\|_{1,\infty} < \frac{1}{|1-\rho_{22}|}$. Therefore, the parameter space property 1 is fulfilled if $|\rho_{11}| < 1$, $|\rho_{22}| < 1$ and $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$.

Since $\|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_1 \wedge \|\rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_\infty < 1$ hold under the proposed parameter space it follows due to Lemma 1, the boundedness of $\mathbf{A}_1$ and $\mathbf{A}_2$ and $\max\{||\mathbf{W}_{ni,nj}||_\varrho$ where $\varrho \in \{1,\infty\}$ and $i,j \in \{1,2\}\} \leq 1$ that parameter space property 2 is also fulfilled.

The parameter space $|\rho_{11}| < 1$, $|\rho_{22}| < 1$ and $\frac{|\rho_{12}|}{|1-\rho_{22}|}\frac{|\rho_{21}|}{|1-\rho_{11}|} < 1$ obviously fulfills the parameter space property 3.

Part B: The following assumptions are being made: $\mathbf{W}_{n1,n1}$, $\mathbf{W}_{n2,n2}$, $\mathbf{W}_{n2,n1}$ and $\mathbf{W}_{n1,n2}$ are row standardized, $\mathbf{W}_{n1,n1}$ and $\mathbf{W}_{n2,n2}$ can be written as $\mathbf{W}_{n1,n1} = \Lambda_{n1,n1}\overline{\mathbf{W}}_{n1,n1}, \mathbf{W}_{n1,n2} = \Lambda_{n1,n2}\overline{\mathbf{W}}_{n1,n2}, \mathbf{W}_{n2,n1} = \Lambda_{n2,n1}\overline{\mathbf{W}}_{n2,n1}$ and $\mathbf{W}_{n2,n2} = \Lambda_{n2,n2}\overline{\mathbf{W}}_{n2,n2}$ where Λ represents the row-standardizing and both $\overline{\mathbf{W}}_{n1,n1}$ and $\overline{\mathbf{W}}_{n2,n2}$ are symmetric, $\left\|\overline{\mathbf{W}}\right\|_\infty, \|\mathbf{W}_{n1,n2}\|_\infty$ $\mathbf{W}'_{n2,n1} = \mathbf{W}_{n1,n2}$ and $\|\mathbf{W}_{n2,n1}\|_\infty < \infty$

It has be shown that $\|\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_\infty < \infty$. Under the assumptions, it holds: $\mathbf{A}_1 =$ $\mathbf{A}_2 =$....

$$\|\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_1 = \sum_{k=0}^{\infty}|\rho_{12}\rho_{21}|^k\left\|(\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1})^k\right\|_1 =$$
$$\left\|\mathbf{I}_{n1} + \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1} + \sum_{k=2}^{\infty}|\rho_{12}\rho_{21}|^k\left(\mathbf{A}_1\left(\prod_{j=1}^{k-1}\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\right)\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\right)\right.$$
$$1 + \frac{|\rho_{12}\rho_{21}|\overline{\kappa}_{11}\overline{\kappa}_{22}\overline{\kappa}_{12}\overline{\kappa}_{21}}{|1-\rho_{22}||1-\rho_{11}|} + \sum_{k=2}^{\infty}\frac{|\rho_{12}\rho_{21}|\overline{\kappa}_{11}\overline{\kappa}_{22}\overline{\kappa}_{12}\overline{\kappa}_{21}}{|1-\rho_{22}||1-\rho_{11}|}(|\rho_{12}\rho_{21}|)^{k-1}\left\|\prod_{j=1}^{k-1}\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\right\|_1$$

since $\left\|\prod_{j=1}^{k-1}\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\mathbf{A}_1\right\|_1 = \left\|\prod_{j=1}^{k-1}\mathbf{A}'_1\mathbf{W}_{n2,n1}\mathbf{A}'_2\mathbf{W}'_{n1,n2}\right\|_\infty =$ $\left\|\prod_{j=1}^{k-1}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\right\|_\infty \leq \left(\frac{1}{|1-\rho_{22}||1-\rho_{11}|}\right)^{k-1}$ it follows $\|\mathbf{I}_{n1} - \rho_{12}\rho_{21}\mathbf{A}_1\mathbf{W}_{n1,n2}\mathbf{A}_2\mathbf{W}_{n2,n1}\|_1 \leq$ $(1 + \overline{\kappa}_{11}\overline{\kappa}_{22}\overline{\kappa}_{12}\overline{\kappa}_{21})\sum_{k=0}^{\infty}\left(\frac{|\rho_{12}\rho_{21}|}{|1-\rho_{22}||1-\rho_{11}|}\right)^k < \infty$ ■

Lemma 4 *If* $\varrho \in \{1,\infty\}$ *and* $\|\mathbf{\Psi}_{n,p}\|_\varrho < 1$ *then* $\left\|(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1}\right\|_\varrho \leq \frac{1}{1-\|\mathbf{\Psi}_{n,p}\|_\varrho} < \infty$

Proof. Since $\|\mathbf{\Psi}_{n,p}\|_\varrho < 1$ we apply the Neumannseries to write $(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1} = \sum_{k=0}^{\infty} \mathbf{\Psi}_{n,p}^k$. Therefore, $\left\|(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1}\right\|_\varrho = \left\|\sum_{k=0}^{\infty} \mathbf{\Psi}_{n,p}^k\right\|_\varrho \leq \sum_{k=0}^{\infty} \left\|\mathbf{\Psi}_{n,p}^k\right\|_\varrho \leq \sum_{k=0}^{\infty} \|\mathbf{\Psi}_{n,p}\|_\varrho^k = \frac{1}{1-\|\mathbf{\Psi}_{n,p}\|_\varrho} < \infty$. The inequality follows due to the triangle inequality (see Horn Johnson) and the second due to the sub- multiplicativity of these matrix norms. ■

Neumannseries[16]**:** *If* $\|\mathbf{\Psi}_n\| < 1$ *for any matrix norm it follows*[17]*:* $(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1} = \sum_{k=0}^{\infty} \Psi_n^k$

Proof. First: Let $\|\mathbf{\Psi}_n\| < 1$ then: $\lim_{k\to\infty} \left\|\mathbf{\Psi}_n^k\right\| \leq \lim_{k\to\infty} \|\mathbf{\Psi}_n\|^k = \mathbf{0}_{n,n}$ (for the first inequality we use the sub- multiplicativity of these matrix norms). Therefore, $\lim_{k\to\infty} \mathbf{\Psi}_n^k = \mathbf{0}_{n,n}$. Second, we show $(\mathbf{I}_n - \mathbf{\Psi}_{n,p})^{-1} = \sum_{k=0}^{\infty} \mathbf{\Psi}_n^k$: $(\mathbf{I}_n - \mathbf{\Psi}_{n,p}) \sum_{k=0}^{\infty} \mathbf{\Psi}_n^k \stackrel{?}{=} \mathbf{I}_n \Leftrightarrow (\mathbf{I}_n - \mathbf{\Psi}_{n,p}) \sum_{k=0}^{\infty} \mathbf{\Psi}_n^k = \lim_{k\to\infty} \mathbf{I}_n - \mathbf{\Psi}_{n,p}^k \stackrel{?}{=} \mathbf{I}_n \Leftrightarrow \mathbf{I}_n \stackrel{?}{=} \mathbf{I}_n$ ■

REFERENCES

[1] **Bronstein, Semendjajew, Musiol, Mühlig.** 2000. *Taschenbuch der Methematik*. Harri Deutsch,Frankfurt am Main

[2] **Elhorst P. J.** 2010. *Applied Spatial Econometrics: Raising the Bar*. Spatial Economic Analysis **5(1)**: p.9-28

[3] **Hamliton J.** 1994. *Time Series Analysis.* Princton University Press, Princton, New Jersy

[4] **Heuser H.** 1995. *Gewöhnliche Differentialgleichunegen.* Vieweg + Teubner, Wiesbaden

[5] **Horn R., Johnsen C.** 1985. *Matrix analysis.* Cambridge University Press, Cambridge, London, New York, Melbourne, Sydney

[16] The Neumann series was partly developed by Carl Neumann (1832-1925) who used it in the context of potential theory (1877). It is a useful "tool" in functional analysis. We point that out, since some authors use the for the "Neumannseries" the incorrect term "Taylorseries".

[17] The series is even more general: Let Ξ be Banach algebra and $x \in \Xi$ where $\|x\| < 1$. Then the series $\sum_{k=0}^{\infty} x^k$ is absolutely convergent, and $\sum_{k=0}^{\infty} x^k = (e-x)^{-1}$ where e denotes the one-element in Ξ. [for more details see for example Heuser 1995]

[6] **Kelejian H., Prucha I.** 1998. *A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.* The Journal of Real Estate Finance and Economics **17**: p.99-121

[7] **Kelejian H., Prucha I.** 2008. *Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances.* CESIFO working paper No. 2448

[8] **Koch M.** 2010. *A research note on asymptotic parameter spaces: A proof for the spatial one forward one behind pattern.* working paper

[9] **Lee L.F.** 2010. *Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances.* Econometric Reviews, **22**: p.307-335.

[10] **Lee L.F., Liu X .** 2010. *Efficient GMM Estimation of High Order Spatial Autoregressive Models with Autoregressive Disturbances.* Econometric Theory **26**: p.187-230

[11] **Lee L.F., Yu J.** 2010. *Estimation of spatial autoregressive panel data models with fixed effects.* Journal of Econometrics **154**: p.165-185

[12] **Pace K., LeSage J and Zhu Shuang** 2010. *Spatial Dependence in Regressors and its Effect on Estimator Performance.* working paper